



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

Bayesian latent class models to estimate diagnostic test accuracies of COVID-19 tests

Hartnack, Sonja ; Eusebi, Paolo ; Kostoulas, Polychronis

Abstract: With great interest we read the article from Dramé et al. 1 “Should RT-PC be considered as gold standard in the diagnosis of Covid-19” questioning the conclusions from Cassaniti et al.2. We agree with the argument that considering RT-PCR as gold standard when evaluating a new test will inevitably lead to biased test accuracy estimates of the new test1.

DOI: <https://doi.org/10.1002/jmv.26405>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-189251>

Journal Article

Accepted Version

Originally published at:

Hartnack, Sonja; Eusebi, Paolo; Kostoulas, Polychronis (2021). Bayesian latent class models to estimate diagnostic test accuracies of COVID-19 tests. *Journal of Medical Virology*, 93(2):639-640.

DOI: <https://doi.org/10.1002/jmv.26405>

Sonja Hartnack ORCID iD: 0000-0002-5757-5708

Paolo Eusebi ORCID iD: 0000-0002-0715-6396

Title

Bayesian latent class models to estimate diagnostic test accuracies of COVID-19 tests

Running title: Bayesian latent class models and test accuracies of COVID-19 tests

Authors: Sonja Hartnack^{1*}, Paolo Eusebi², Polychronis Kostoulas³

¹Section of Epidemiology, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

²Health Planning Service, Regional Health Authority of Umbria, Perugia, Italy

³School of Health Sciences, University of Thessaly, Karditsa, Greece

*Corresponding author:

Sonja Hartnack

Section of Epidemiology

Vetsuisse Faculty Zurich

University of Zurich

Winterthurerstr. 270

8057 Zurich

Switzerland

Email: sonja.hartnack@access.uzh.ch

Phone: +41 44 635 90 51

To the Editor,

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jmv.26405.

This article is protected by copyright. All rights reserved.

With great interest we read the article from Dramé et al.¹ “Should RT-PCR be considered as gold standard in the diagnosis of Covid-19” questioning the conclusions from Cassaniti et al.². We agree with the argument that considering RT-PCR as gold standard when evaluating a new test will inevitably lead to biased test accuracy estimates of the new test¹. However, we disagree that it is impossible to “demonstrate that a new test might be better than the old-test” and propose using Bayesian latent class models (BLCMs) to estimate the sensitivity and specificity in the absence of a gold standard. “Latent” means here that the true disease/infection status is not observed, but can be estimated from the data. BLCMs have been applied widely, reporting guidelines have been developed³, and currently an EU-funded initiative aims to expand their application in the evaluation of diagnostic tests (www.harmony-net.eu).

For show-casing BLCMs we used the data from Cassaniti et al.², although the keys to validity in diagnostic test studies are not respected, i.e. “a consecutive series of patients suspected (but not known) to have the target disorder”⁴ - which precludes the generalisability of the findings. We considered a model with three tests for the diagnosis of COVID-19 in one population (110 patients), and extracted the following data of cross-classified test results (RT-PCR, IgG and IgM)²: +++ 28, ++- 1, +-+ 3, +-- 36, -++ 0, -+- 0, --+ 1, --- 41. We aim to estimate the prevalence, the sensitivities and the specificities of three tests, and potential conditional dependencies (correlations) between tests. Since the maximum number of estimable parameters for three tests in one population is seven⁵, this model is not identifiable. Therefore, in a Bayesian framework, we rely on prior information for the specificities of all three tests similar to a previous analysis with two tests and four populations⁶. We assume that they are close to 100%. Further technical details about the model and Markov Chain Monte Carlo simulation, including the code to reproduce the results, are available on

(<https://github.com/shartn/BLCM-COVID19>). In Table 1 the posteriors resulting from a model with conditional dependency between the IgG and IgM test sensitivities are displayed.

In contrast to the previous analysis ², which used RT-PCR as a gold standard, implicitly assuming 100% sensitivity and specificity, the RT-PCR sensitivity based on our BLCM analysis is lower, while the specificity of all three tests is close to 100%. The estimated sensitivity of the combined IgM/IgG in the study from Cassaniti et al. ² was 18.4%, being outside the 95% probability interval of our BLCM estimates for both the IgM and IgG. Although, based on our BLCM analysis, we cannot state “that the new test might be better” ¹, we are confident that BLCMs allow estimation of sensitivities and specificities in an unbiased way without needing a gold standard. Since this data set - with the exception of the 50 emergency patients - does not comply with the key aspects of validity ⁴, it is not possible to generalise the results of our BLCM beyond the study population of the 110 patients.

Despite enormous COVID-19 research activities, the application of BLCMs and suitable data sets are virtually absent. For the preparedness of future pandemics, we suggest to include the design of appropriate diagnostic test accuracy studies, which will also be suitable for BLCMs.

Table 1 - Posterior medians and 95% probability intervals for the sensitivities and specificities of three tests, prevalence and the correlation between the sensitivities of the IgM and the IgG test resulting from BLCM analysis

Parameter	Posterior median (95% Probability interval)
Prevalence	87.9 (72.2;99.0)
RT-PCR Sensitivity	55.2 (49.7;65.6)
RT-PCR Specificity	99.0 (97.7;99.7)
IgG Sensitivity	31.3 (21.9;42.4)
IgG Specificity	98.2 (94.8;99.6)
IgM Sensitivity	33.2 (23.8;44.1)
IgM Specificity	98.1 (94.4;99.6)
Correlation IgG/IgM	8.6 (6.3;10.4)

References

1. Dramé M, Teguo MT, Proye E, et al. Should RT-PCR be considered a gold standard in the diagnosis of Covid-19? *J Med Virol*. 2020:0-2. doi:10.1002/jmv.25996
2. Cassaniti I, Novazzi F, Giardina F, et al. Performance of VivaDiag COVID-19 IgM/IgG Rapid Test is inadequate for diagnosis of COVID-19 in acute patients referring to emergency room department. *J Med Virol*. 2020:1-4. doi:10.1002/jmv.25800
3. Kostoulas P, Nielsen SS, Branscum AJ, et al. STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. *Prev Vet Med*. 2017;138:37-47. doi:10.1016/j.prevetmed.2017.01.006
4. Sackett DL, Haynes RB. The architecture of diagnostic research. *Bmj*. 2002;324:539-541. doi:10.1136/bmj.324.7337.602
5. Berkvens D, Speybroeck N, Praet N, Adel A, Lesaffre E. Estimating disease prevalence in a Bayesian framework using probabilistic constraints. *Epidemiology*. 2006;17(2):145-153. doi:10.1097/01.ede.0000198422.64801.8d
6. Kostoulas P, Eusebi P, Hartnack S. Diagnostic accuracy estimates for COVID-19 RT-PCR and Lateral flow immunoassay tests with Bayesian latent class models. :1-11. doi:10.21203/rs.3.rs-33243/v1